

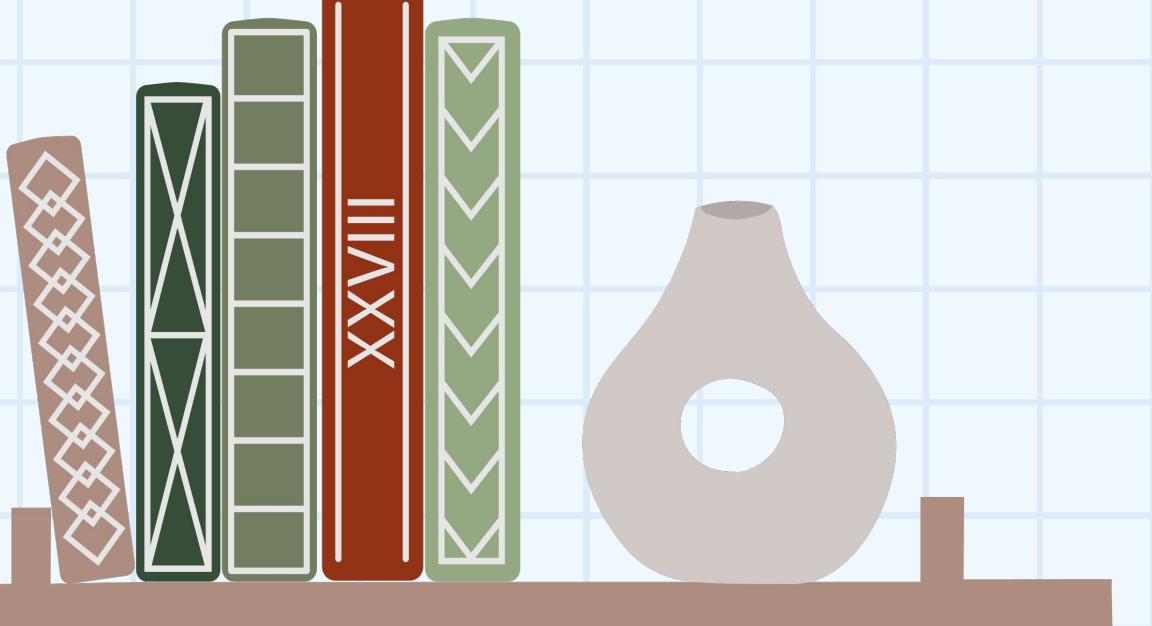
Powered by
Futurense



BS./BSC.IN

Applied AI and Data Science

Basics of Data Analytics



Let's dive into and learn:



- 1 Data Warehouses
- 2 Data Marts
- 3 Data Lakes



Data Warehouse



Powered by

Futuraense

- Data warehouse refers to a large central repository where data from multiple sources are stored
- Consolidates data coming from disparate sources
- Uses the Extract, Transform, Load process



Data Warehouse

- Works as a multipurpose storage facility of data storage for different purposes
- Data comes to data warehouses in a structured format
- Data has already been cleaned and organized
- Data is analysis ready
- Stores recent and historic data
- Typically opt for a data warehouse when we have large amounts of data



Data Marts

- Data Marts are a particular section of a Data warehouse built for a particular purpose
- Could a particular business or a particular community of users
- Contains data most relevant to set of stakeholders
- E.g. a data mart could be the customer transactions data for a particular business. Used to make sales projections or marketing campaign

Data Marts



Powered by

Futurense

- Data Marts have analytical capabilities for a restricted part of the data warehouse
- Has isolate security and isolated performance
- Usually used for business specific reporting and analytics

Data Lake



- Data Lakes are storage repositories that can store large amounts of data irrespective of structure
- Store large amounts of structured, unstructured and semi-structured data
- Usually in original or native format
- Each data element has a unique identifier
- With tags for each segment for identification and further use

Data Lake



Powered by

Futurense

- Opt for a data lake if we process or have access to large volumes of data on an ongoing basis
- Don't want to commit to or be restricted by some specific or pre-determined use case
- Contain all of the source data across all types and file formats
- Needed for predictive and advanced analytics

Big Picture



Powered by

Futurense

- Data repositories exist as a data storage and holding platform
- Make reporting and analytics more streamlined and credible
- Helps bring all the data together in one place

Recap



Powered by
Futurense

Data Warehouses, Data Marts and Data Lakes



Powered by



Futurense

Thank you

